COURS

Loi binomiale — Épreuves indépendantes et Bernoulli

Tle

Programme de terminale générale, enseignement de spécialité — BO spécial n°8 du 25 juillet 2019

Lancer dix fois une pièce, répondre au hasard à un QCM, vendre plus de billets que de places : chaque fois, on répète la même expérience à deux issues, sans que les essais s'influencent. **Dans un schéma de Bernoulli de paramètres n et p, le nombre X de succès suit la loi binomiale $\mathcal{B}(n,p)$: pour tout entier k de 0 à n, $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$.** Le coefficient binomial compte les chemins à k succès, la puissance donne la probabilité de chacun.

1 Succession d'épreuves, arbres et indépendance

DÉFINITION

On réalise successivement n épreuves, d'univers $\Omega_1,\dots,\Omega_n$, **sans que le résultat de l'une influe sur les suivantes**. Une issue est un n-uplet $(x_1,\dots,x_n)$ du produit cartésien $\Omega_1\times\cdots\times\Omega_n$, c'est-à-dire un **chemin** de l'arbre. Dans ce **modèle**, la probabilité d'une issue est le **produit** des probabilités de ses composantes, $P\big((x_1,\dots,x_n)\big)=P(x_1)\times\cdots\times P(x_n)$: dans l'arbre, chaque épreuve porte les **mêmes pondérations**, quel que soit le chemin suivi.

PROPRIÉTÉ

Rappel de première : l'arbre pondéré. Pour des épreuves **quelconques**, une branche du deuxième niveau porte une **probabilité conditionnelle** $P_A(B)$; les branches issues d'un nœud ont pour somme 1. Un chemin a pour probabilité le produit de ses branches, $P(A\cap B)=P(A)\times P_A(B)$. Si $A_1,\dots,A_m$ partitionnent l'univers, **probabilités totales** : $P(B)=P(A_1\cap B)+\cdots+P(A_m\cap B)$. A et B sont **indépendants** si $P(A\cap B)=P(A)\times P(B)$.

MODÉLISER UNE SUCCESSION D'ÉPREUVES

1. **Découper l'expérience en épreuves** et décider si elles sont **indépendantes** : lancers successifs, tirages **avec remise**… Sans remise, la composition de l'urne change : les pondérations du deuxième niveau sont des probabilités conditionnelles.
   l'indépendance fait partie du modèle : l'énoncé la donne ou la situation l'impose ; on ne la démontre pas.
2. **Construire l'arbre**, multiplier le long d'un chemin, additionner les chemins qui réalisent l'événement (probabilités totales).

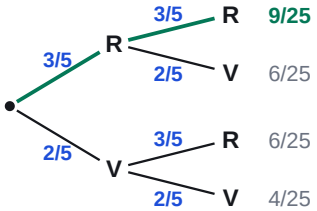
EXEMPLE

Une urne contient 3 boules rouges (R) et 2 vertes (V). On tire deux boules l'une après l'autre. Probabilité d'obtenir deux rouges, puis que la seconde soit rouge, avec et sans remise ?

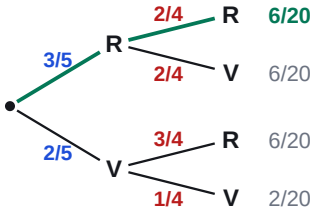
1. **Avec remise**, tirages indépendants : chaque branche R porte $\frac{3}{5}$, et $P(R_2)=\frac{3}{5}$.
   $P(R_1\cap R_2)=\frac{3}{5}\times\frac{3}{5}=\frac{9}{25}$
2. **Sans remise**, après une rouge il reste 2 rouges sur 4 boules :
   $P(R_1\cap R_2)=\frac{3}{5}\times\frac{2}{4}=\frac{6}{20}=\frac{3}{10}$
3. Probabilités totales, selon la première boule :
   $P(R_2)=\frac{3}{5}\times\frac{2}{4}+\frac{2}{5}\times\frac{3}{4}=\frac{12}{20}=\frac{3}{5}$

$\frac{9}{25}$ avec remise, $\frac{3}{10}$ sans remise, où $P_{R_1}(R_2)=\frac{1}{2}\neq P(R_2)$: les tirages ne sont pas indépendants.

avec remise



sans remise



avec remise : Ω₁ × Ω₂

1er \ 2e	R (3/5)	V (2/5)
R (3/5)	3/5 × 3/5 **9/25**	3/5 × 2/5 **6/25**
V (2/5)	2/5 × 3/5 **6/25**	2/5 × 2/5 **4/25**

Avec remise, les deux niveaux portent les mêmes pondérations. Sans remise, celles du second niveau (en rouge) dépendent du premier tirage. En vert, le chemin « deux rouges ».

PIÈGE

Sans remise, écrire $P(R_1 \cap R_2) = \frac{3}{5} \times \frac{3}{5}$. — le produit des probabilités des composantes ne vaut que pour des épreuves **indépendantes** ; ici la seconde pondération est $P_{R_1}(R_2) = \frac{2}{4}$.

Ce que ça coûte : le modèle entier, donc toutes les probabilités qui en découlent

Le réflexe : avant de multiplier, se demander si la première épreuve change les conditions de la seconde.

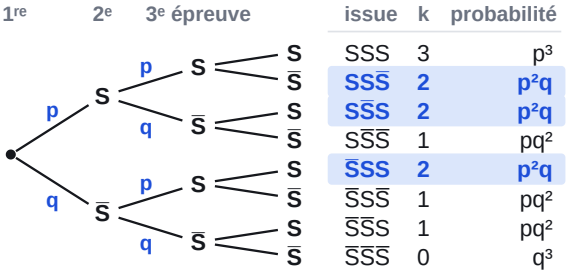
2 Épreuve, loi et schéma de Bernoulli

DÉFINITION

Une **épreuve de Bernoulli** est une épreuve à **deux issues**, le **succès** S et l'**échec** $\overline{S}$, avec $P(S) = p$ et $P(\overline{S}) = 1 - p$. La variable aléatoire X qui vaut 1 en cas de succès et 0 en cas d'échec suit la **loi de Bernoulli de paramètre** p. Calcul direct : $E(X) = 0 \times (1-p) + 1 \times p = p$ et, comme $X^2 = X$, $V(X) = E(X^2) - E(X)^2 = p - p^2 = p(1-p)$.

DÉFINITION

Un **schéma de Bernoulli de paramètres n et p** est la répétition de n épreuves de Bernoulli **identiques et indépendantes**, de même probabilité de succès p. Son univers est $\{S, \overline{S}\}^n$, ou $\{0,1\}^n$: une issue est un mot de n lettres, un chemin de l'arbre. Par le modèle des épreuves indépendantes, **une issue à k succès, donc à $n-k$ échecs, a pour probabilité $p^k(1-p)^{n-k}$**, quel que soit l'ordre des lettres.



issue	k	probabilité
SSS	3	p^3
SS$\overline{S}$	2	p^2q
S$\overline{S}$S	2	p^2q
S$\overline{S}\overline{S}$	1	pq^2
$\overline{S}$SS	2	p^2q
$\overline{S}$S$\overline{S}$	1	pq^2
$\overline{S}\overline{S}$S	1	pq^2
$\overline{S}\overline{S}\overline{S}$	0	q^3

q = 1 − p

regrouper les issues selon k

k	chemins	P(X = k)
0	1	q^3
1	3	$3pq^2$
2	3	$3p^2q$
3	1	p^3

1, 3, 3, 1 : les coefficients binomiaux
« 3 parmi k », ligne 3 du triangle de Pascal

Schéma de Bernoulli pour $n = 3$: chaque branche haute porte p, chaque branche basse $q = 1 - p$. Les trois issues bleues ont 2 succès et la même probabilité p^2q.

RECONNAÎTRE UN SCHÉMA DE BERNOULLI

1. **Nommer l'épreuve répétée et le succès** (« obtenir 6 », « la pièce est défectueuse »), puis lire p.
2. **Vérifier les conditions** : deux issues ; les **mêmes** conditions à chaque essai, donc p constant ; des essais **indépendants**. Un prélèvement dans une très grande population est parfois **assimilé** à un tirage avec remise : c'est alors l'énoncé qui le dit.
3. **Compter les répétitions** : n est fixé à l'avance.
   « lancer un dé jusqu'à obtenir 6 » n'est pas un schéma de Bernoulli : le nombre d'essais n'est pas fixé.

EXEMPLE

On lance trois fois un dé équilibré ; le succès est « obtenir 6 ». Calculer la probabilité d'obtenir exactement deux 6.

1. Schéma de Bernoulli de paramètres $n = 3$ et $p = \frac{1}{6}$. Trois chemins ont deux succès, $SS\overline{S}$, $S\overline{S}S$ et $\overline{S}SS$, chacun de probabilité $p^2(1-p)$:
   $3 \times \left(\frac{1}{6}\right)^2 \times \frac{5}{6} = \frac{15}{216} = \frac{5}{72}$

la probabilité d'obtenir exactement deux 6 est $\frac{5}{72} \approx 0{,}069$.

PIÈGE

Donner $\left(\frac{1}{6}\right)^2 \times \frac{5}{6}$ comme probabilité d'obtenir deux 6 en trois lancers. — c'est la probabilité d'**un** chemin, par exemple $SS\overline{S}$; l'événement en compte trois.

Ce que ça coûte : le résultat, faux d'un facteur égal au nombre de chemins

Le réflexe : compter les chemins qui réalisent l'événement : c'est le rôle du coefficient binomial.

3 La loi binomiale

DÉFINITION

Pour $0 \leqslant k \leqslant n$, $\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$ est le nombre de parties à k éléments d'un ensemble à n éléments, donc le nombre de **mots** de n lettres contenant exactement k lettres S, ou de **chemins** à k succès dans l'arbre d'un schéma de Bernoulli. On a $\binom{n}{0} = \binom{n}{n} = 1$, $\binom{n}{1} = n$, $\binom{n}{2} = \frac{n(n-1)}{2}$, la symétrie $\binom{n}{n-k} = \binom{n}{k}$ et, pour $k \leqslant n-1$, la relation de Pascal $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$. Tout est établi dans la fiche « Dénombrement ».

DÉFINITION

Dans un schéma de Bernoulli de paramètres n et p, la variable aléatoire X égale au **nombre de succès** suit la **loi binomiale de paramètres n et p**, notée $\mathcal{B}(n,p)$; elle prend les valeurs $0, 1, \ldots, n$. Si X_i vaut 1 quand la i-ième épreuve est un succès et 0 sinon, $X = X_1 + X_2 + \cdots + X_n$: **une somme de n variables de Bernoulli indépendantes, de même paramètre p.** On en déduit $E(X) = np$ et $V(X) = np(1-p)$: voir la fiche « Sommes de variables aléatoires ».

PROPRIÉTÉ

Probabilité de k succès (démontrée). Si X suit $\mathcal{B}(n,p)$, pour tout entier k tel que $0 \leqslant k \leqslant n$, on a la formule ci-dessous.
Démonstration. L'événement $\{X = k\}$ est formé des issues comportant exactement k succès. Par indépendance, chacune a pour probabilité le produit de k facteurs p et de $n-k$ facteurs $1-p$, soit $p^k(1-p)^{n-k}$. Une telle issue est déterminée par les **rangs** de ses k succès, c'est-à-dire par une partie à k éléments de $\{1, \ldots, n\}$: il y en a $\binom{n}{k}$. On additionne donc $\binom{n}{k}$ fois la même probabilité.

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

CALCULER AVEC UNE LOI BINOMIALE

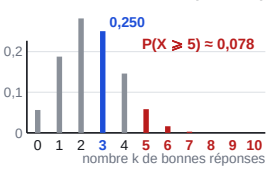
1. **Justifier le modèle** : « on répète n fois, de façon indépendante, une même épreuve de Bernoulli de succès … de probabilité p ; X compte les succès, donc suit $\mathcal{B}(n,p)$ ».
2. **Calculer** $P(X = k)$ par la formule ou la calculatrice (probabilité **ponctuelle**), et $P(X \leqslant k)$ par la probabilité **cumulée**, somme des $P(X = i)$ pour i de 0 à k.
3. **Se ramener au cumul** : $P(X \geqslant k) = 1 - P(X \leqslant k-1)$, $P(k \leqslant X \leqslant k') = P(X \leqslant k') - P(X \leqslant k-1)$; au moins un succès : $P(X \geqslant 1) = 1 - (1-p)^n$.

EXEMPLE

Un QCM compte 10 questions, chacune à quatre propositions dont une seule juste. Un candidat répond au hasard, indépendamment d'une question à l'autre. Calculer la probabilité qu'il ait exactement 3 bonnes réponses, puis au moins 5.

1. Chaque question est une épreuve de Bernoulli (succès : « bonne réponse », $p = 0{,}25$), répétée 10 fois de façon indépendante : X suit $\mathcal{B}(10\,;0{,}25)$. Avec $\binom{10}{3} = 120$:
   $P(X = 3) = 120 \times 0{,}25^3 \times 0{,}75^7 \approx 0{,}2503$
2. « Au moins 5 » est le contraire de « au plus 4 » ; la calculatrice donne $P(X \leqslant 4) \approx 0{,}9219$:
   $P(X \geqslant 5) = 1 - P(X \leqslant 4) \approx 1 - 0{,}9219 = 0{,}0781$

$\underline{P(X = 3) \approx 0{,}25 \text{ et } P(X \geqslant 5) \approx 0{,}078.}$



En bleu, $P(X = 3)$; en rouge, les bâtons dont la somme est $P(X \geqslant 5)$.

PIÈGE

Calculer $P(X \geqslant 5)$ par $1 - P(X \leqslant 5)$. — le contraire de $\{X \geqslant 5\}$ est $\{X \leqslant 4\}$: X ne prend que des valeurs entières, et la valeur 5 appartient à l'événement.

Ce que ça coûte : le résultat, et ceux des seuils qui en dépendent

Le réflexe : écrire l'événement contraire en toutes lettres avant d'utiliser le cumul.

4 Seuils, intervalles et algorithmes

RÉSOUDRE UN PROBLÈME DE SEUIL OU D'INTERVALLE

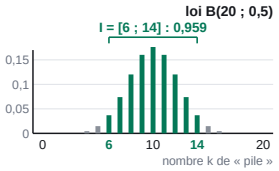
1. **Au moins un succès** : le plus petit n tel que $1-(1-p)^n \geqslant s$ vérifie $(1-p)^n \leqslant 1-s$; passer au logarithme népérien, en **changeant le sens** quand on divise par $\ln(1-p) < 0$, ou tabuler $(1-p)^n$. Pour $p = 0{,}1$ et $s = 0{,}95$: $0{,}9^n \leqslant 0{,}05 \iff n \geqslant \dfrac{\ln 0{,}05}{\ln 0{,}9} \approx 28{,}4$, donc $n = 29$.
2. **Seuil sur k** : pour le plus petit k tel que $P(X > k) \leqslant \alpha$, calculer $P(X \leqslant k)$ pour $k = 0, 1, 2, \ldots$ jusqu'à $1 - P(X \leqslant k) \leqslant \alpha$: $J =]k\,;n]$ vérifie $P(X \in J) \leqslant \alpha$.
3. **Intervalle** : pour $P(X \in I) \geqslant 1-\alpha$, retirer de part et d'autre des valeurs peu probables, puis **vérifier** $P(a \leqslant X \leqslant b) = P(X \leqslant b) - P(X \leqslant a-1) \geqslant 1-\alpha$: X tombe hors de I avec une probabilité au plus α.

EXEMPLE

On lance 20 fois une pièce équilibrée ; X est le nombre de « pile ». Déterminer le plus petit intervalle I centré en 10 tel que $P(X \in I) \geqslant 0{,}95$.

1. X suit $\mathcal{B}(20\,;0{,}5)$. La calculatrice donne $P(X \leqslant 5) \approx 0{,}0207$, $P(X \leqslant 6) \approx 0{,}0577$, $P(X \leqslant 13) \approx 0{,}9423$ et $P(X \leqslant 14) \approx 0{,}9793$. Pour $I = [6\,;14]$: $P(6 \leqslant X \leqslant 14) = P(X \leqslant 14) - P(X \leqslant 5) \approx 0{,}9793 - 0{,}0207 = 0{,}9586$
2. Pour $[7\,;13]$, plus étroit, $0{,}9423 - 0{,}0577 = 0{,}8846 < 0{,}95$.

$I = [6\,;14]$, avec $P(X \in I) \approx 0{,}96$.



En vert, de 6 à 14 : environ 0,959 ; les deux queues grises : environ 0,041.

calculer : loi binomiale et surréservation

```
from math import comb
def proba(n, p, k):        # P(X = k)
    return comb(n, k) * p**k * (1 - p)**(n - k)
def seuil(n, p, alpha):
    k = 0
    F = proba(n, p, 0)     # F = P(X <= k)
    while 1 - F > alpha:   # P(X > k) > alpha
        k = k + 1
        F = F + proba(n, p, k)
    return k
seuil(105, 0.92, 0.05)     # renvoie 101
```

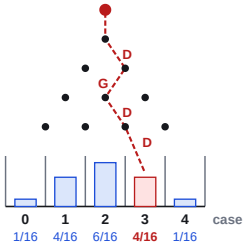
simuler : Bernoulli, binomiale, échantillon

```
from random import random
def bernoulli(p):          # 1 ou 0
    if random() < p:
        return 1
    return 0
def binomiale(n, p):       # nombre de succès
    s = 0
    for i in range(n):
        s = s + bernoulli(p)
    return s
def echantillon(N, n, p): # N valeurs de X
    return [binomiale(n, p) for i in range(N)]
```

`seuil` cumule les $P(X = k)$ jusqu'à $P(X > k) \leqslant \alpha$; `echantillon` construit sa liste en compréhension.

SIMULER UN SCHÉMA DE BERNOULLI EN PYTHON

1. **Simuler** : `bernoulli(p)` renvoie 1 si `random()`, tiré dans $[0\,;1[$, est inférieur à p, donc avec la probabilité p ; `binomiale(n, p)` additionne n appels indépendants et simule $\mathcal{B}(n,p)$; `echantillon(N, n, p)` renvoie un **échantillon** de taille N.
2. **La planche de Galton** : à chacun des n clous, la bille part à droite avec la probabilité 0,5, indépendamment ; sa case, le nombre de rebonds à droite, suit $\mathcal{B}(n\,;0{,}5)$, que simule `binomiale(n, 0.5)`.



Le chemin DGDD mène à la case 3 ; $\binom{4}{k}$ chemins sur 16 mènent à la case k.

À VÉRIFIER

- Schéma de Bernoulli justifié (succès nommé, p, n répétitions **identiques et indépendantes**) ; sinon, retour à l'arbre ?
- Coefficient $\binom{n}{k}$ devant $p^k(1-p)^{n-k}$? Contraire de $\{X \geqslant k\}$ écrit $\{X \leqslant k-1\}$?
- Seuil : sens changé en divisant par $\ln(1-p) < 0$, entier conclu par excès, phrase en contexte ?

À RETENIR

- Épreuves **indépendantes** : la probabilité d'une issue est le **produit** des probabilités de ses composantes.
- Bernoulli : $P(X=1) = p$, $P(X=0) = 1-p$, $E(X) = p$, $V(X) = p(1-p)$.
- Schéma de Bernoulli : n épreuves **identiques et indépendantes** ; un chemin à k succès a pour probabilité $p^k(1-p)^{n-k}$.
- X suit $\mathcal{B}(n,p)$: $P(X=k) = \binom{n}{k} p^k(1-p)^{n-k}$, pour $0 \leqslant k \leqslant n$.
- $P(X \geqslant k) = 1 - P(X \leqslant k-1)$; au moins un succès : $1-(1-p)^n$.